# The Billion-Transistor Budget: A Different Kind of Real Estate Development

**John H. Crawford**
Intel Fellow
Enterprise Platform Group
Intel Corporation

October 15, 2002

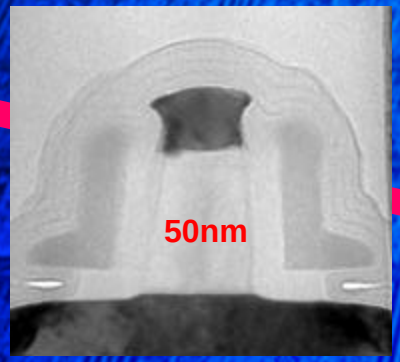



http://www.intel.com/pressroom/kits/events/mpf_2002/index.htm

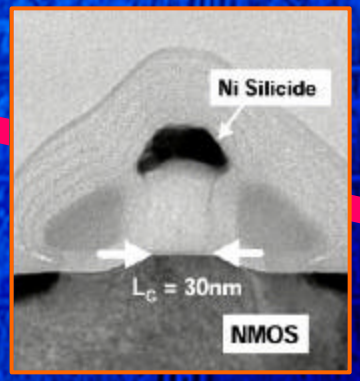
# When Does Moore's Law Get Us to One Billion Transistors?

intel

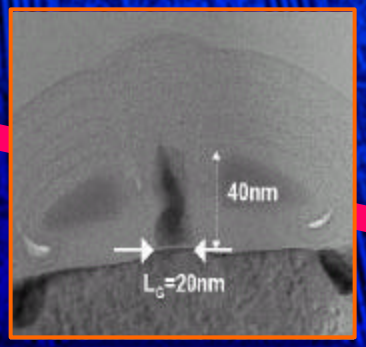
# Nanotechnology Advancements



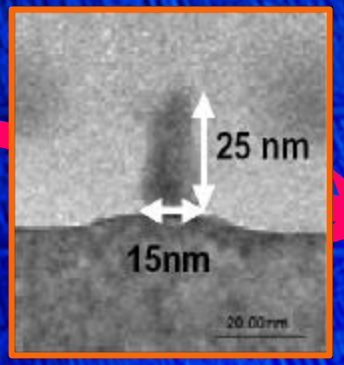

90nm Node
Lgate = 50nm
Production - 2003



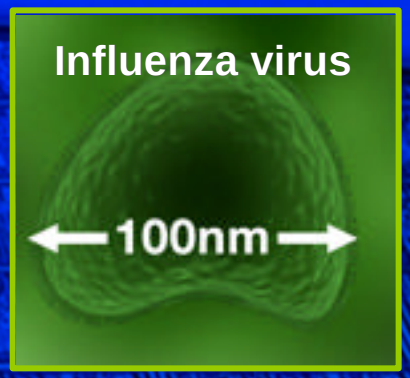

65nm Node
Lgate = 30nm
Production - 2005




45nm Node
Lgate = 20nm
Production - 2007



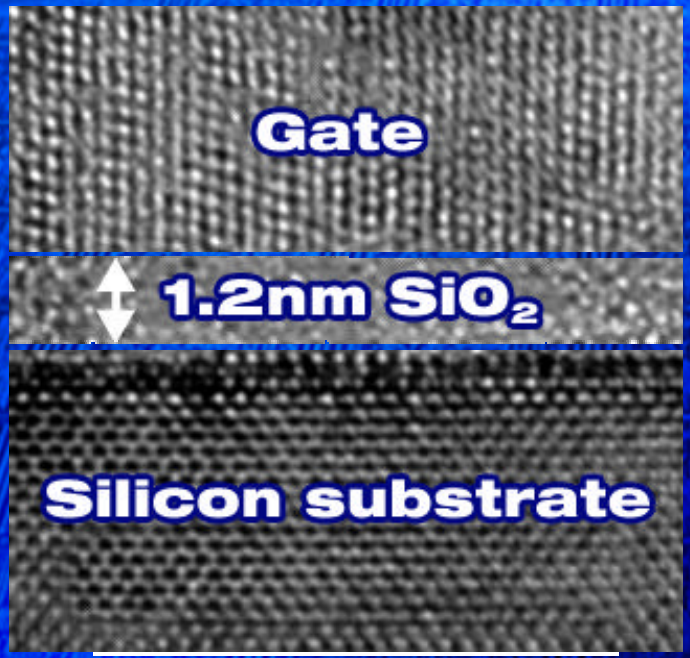

30nm Node
Lgate = 15nm
Production - 2009

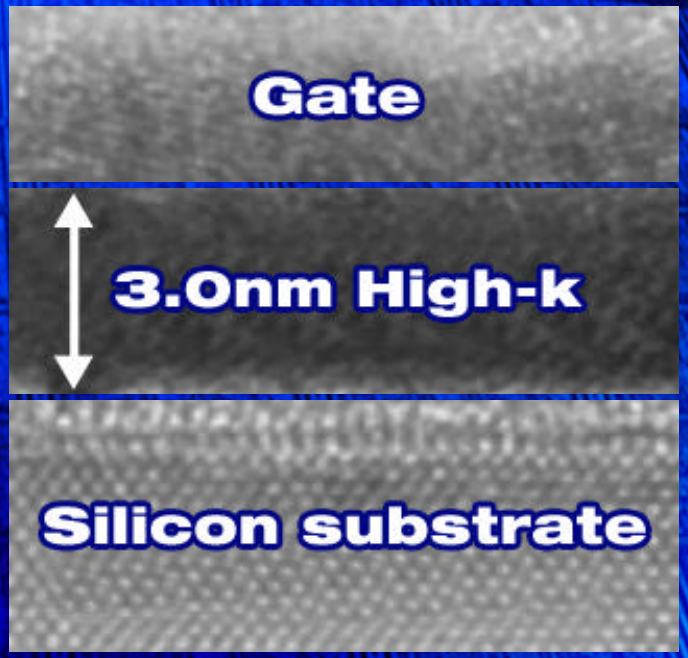



**Process Advancements Fulfill Moore's Law**

Source: Intel

# Nanotechnology Gate Dielectrics



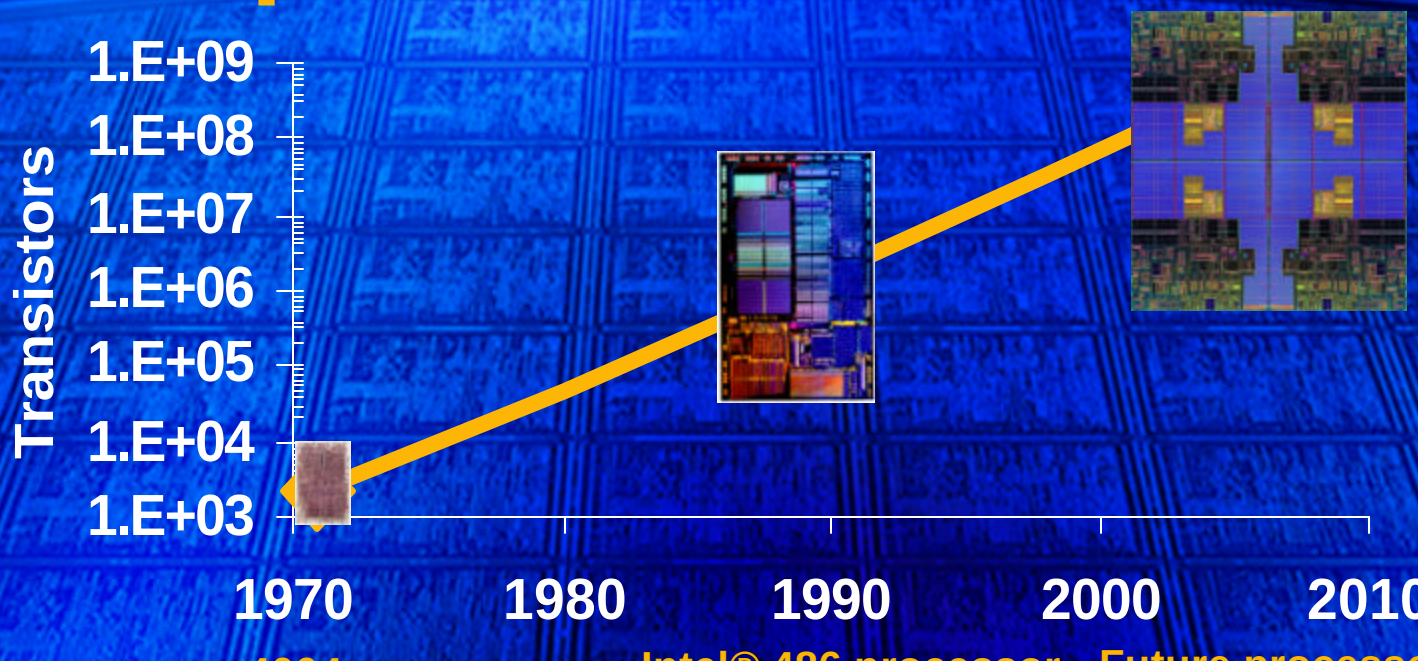




Source: Intel

| | 90nm process | Experimental high-k |
|---|---|---|
| Capacitance | 1X | 1.6X |
| Leakage | 1X | < 0.01X |

**New Materials Support Continued Scaling**

# Microprocessor Evolution




| | 4004 | Intel® 486 processor | Future processor |
|---|---|---|---|
| Date | 1971 | 1989 (18 yrs) | 2007 (+18 yrs) |
| Transistor Count | 2300 transistors | 1.2M (500 X) | 1000M (800 X) |
| Mfg. Process | 10um process | 1.0 um (1/10 X) | 65nm (1/15 X) |
| Wafer Area | 50mm wafer | 100-150mm (4-9X area) | 300mm (4-9 X area) |
| Die Size | 12mm2 | 174mm2 (14 X) | 400mm$^2$ (2.3 X) |
| Core Frequency | 108 kHz | 25 MHz (250 X) | 6 GHz (240 X) |

**Moore's Law delivers 1 Billion Transistors circa 2007**

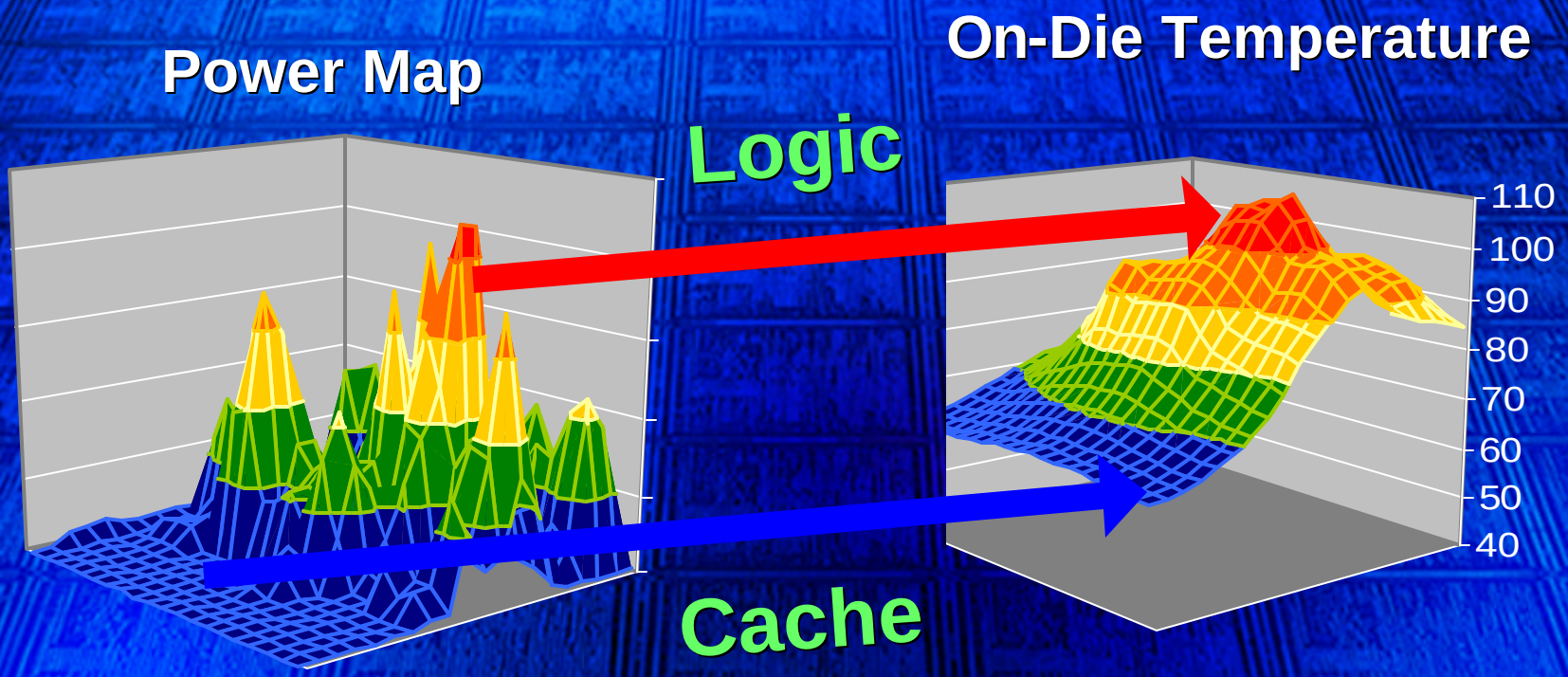
“If the automobile industry advanced as rapidly as the semiconductor industry, a Rolls Royce would get a million miles per gallon and it would be cheaper to throw it away than to park it.”

Gordon Moore,
Intel Corporation

# Design Challenges

- Design Complexity
  - Productivity Tools and Methods Advance
    - …But at slower rate than Moore's Law
- Visibility for Test & Debug
  - Pin Bandwidth/Transistor continues to decline
  - Shrinking dimensions, increasing speeds, …
- Power
  - Power Delivery – di/dt of Amps/nano-second
  - Thermals: Overall power and thermal density

# Power Density




**Logic runs hot - Cache runs cool**

# What Do We Do With a Billion Transistors?

# Market Trends

## Market Segmentation

| | | |
|---|---|---|
| **Itanium® Processor** *Enterprise* | **Banias** *Mobile* | **Intel® IXA2800** *Network Processor* |
| **Intel® Xeon™ Processor** *Server* | **Pentium® 4 Processor** *Desktop* | **Intel® XScale™ Core** *Handheld* |

## Convergence of Computing and Communications

# Technology Trends

## More Performance

- Multi-Core
- Larger Cache
- Thread-Level Parallelism

## More Than Performance

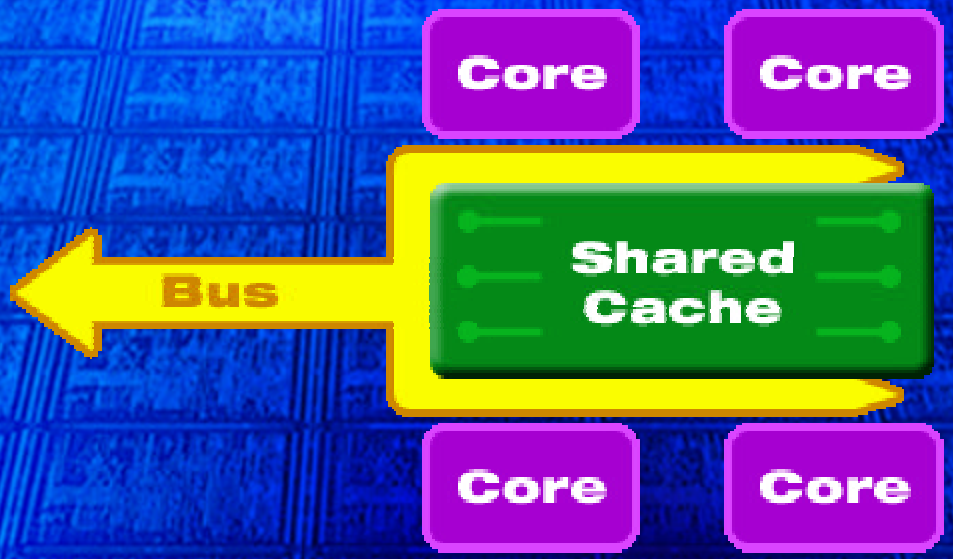
- High Reliability
- Managed Run-Time Environments
- Safer Computing

## Balancing Performance and Power

- Power Management
- Design Tradeoffs
- Die Size

intel

# Multi-Core Enterprise Processor

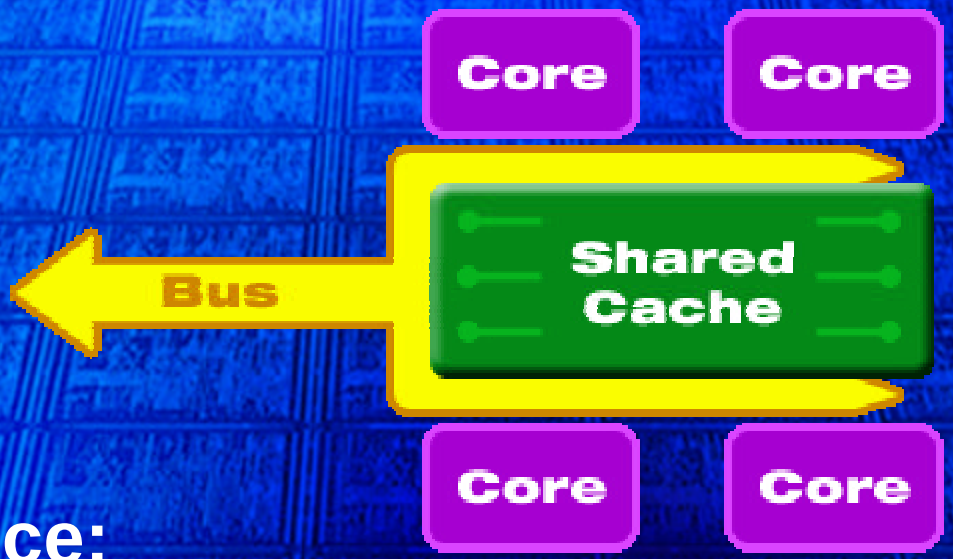



4 Processor system on a chip, Integrating:

- 4 Itanium 2 processor Cores ~120 M transistors
- Shared Cache 12-16 MB 700-950 M transistors
- Leaf interconnect

**1B Transistors Can and Will be Used**

# Multi-Core Enterprise Processor

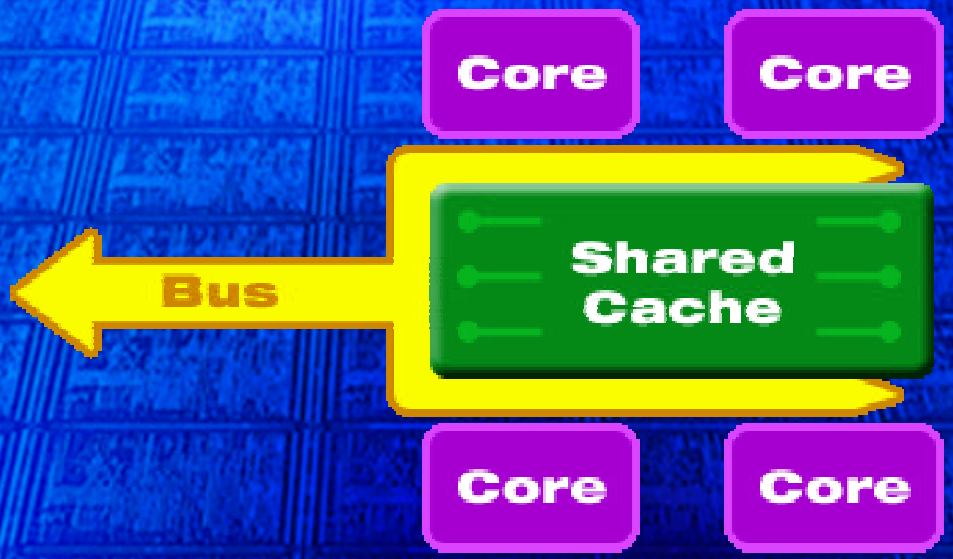



## Performance:

- CMP scales well
- Processor/memory gap widens (50% annual growth vs. 7%)
  - Cache couples high-speed core to slower memory
  - Cache is large, regular structure with low power density
- Shared cache leverages beneficial interference

intel

# Multi-Core Enterprise Processor



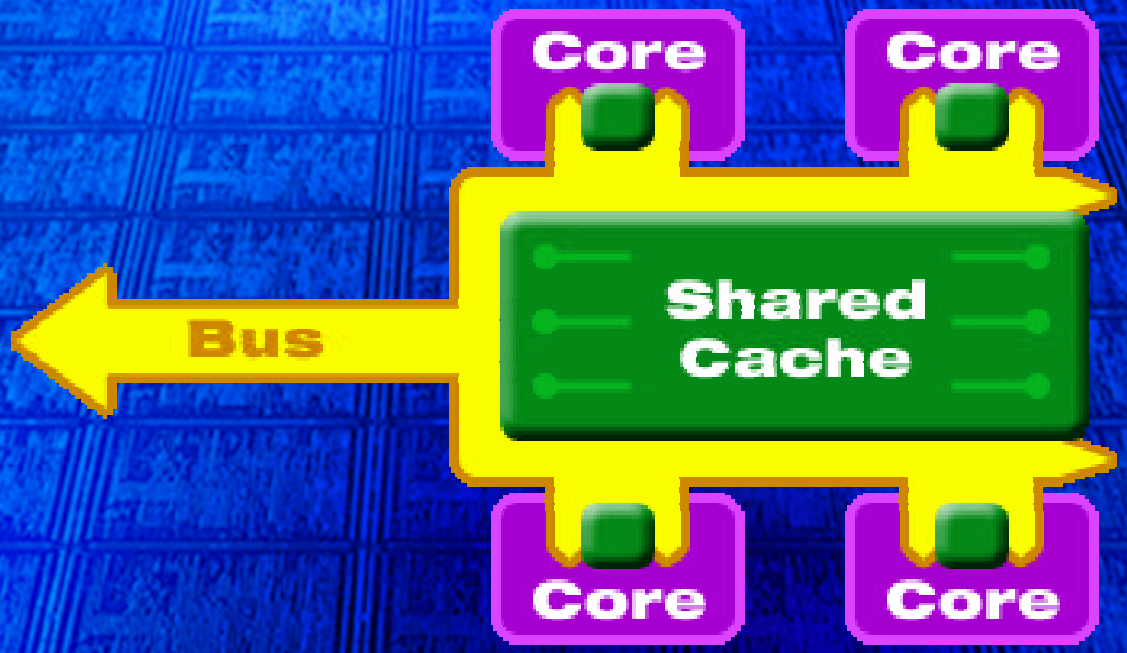

**Power Addressed by:**

- Average out peaks among cores
- Spread hot spots in cores apart
- Cache has low power density
- Power Management techniques

**Complexity Reduced by:**

- Large Cache
- Repeat Cores

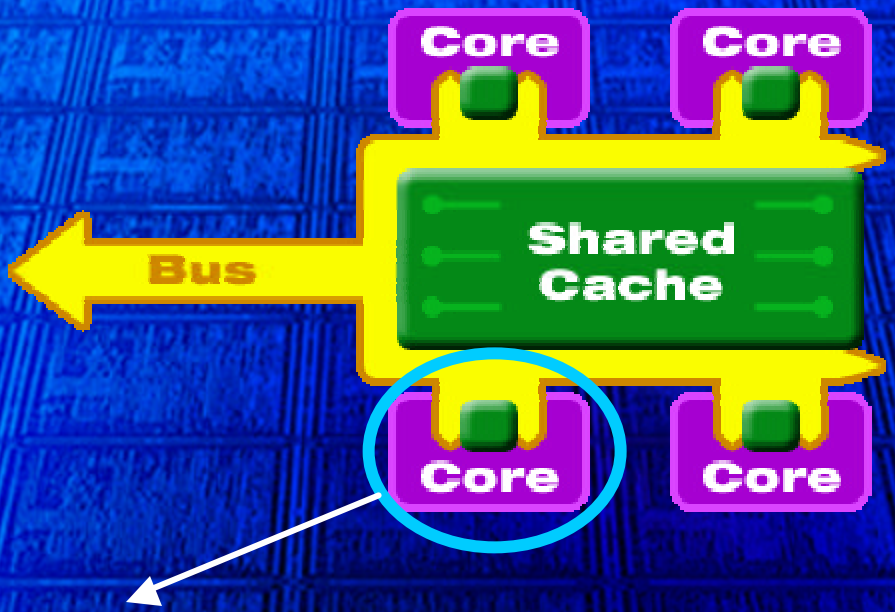
# Multi-Core: On-Chip Interconnect




Internal Interconnect Network

- Integrates “leaves” of system Net
- Optimize for on-chip, fixed topology, multi-level cache control

# Thread-Level Parallelism: Hyper-Threading Technology

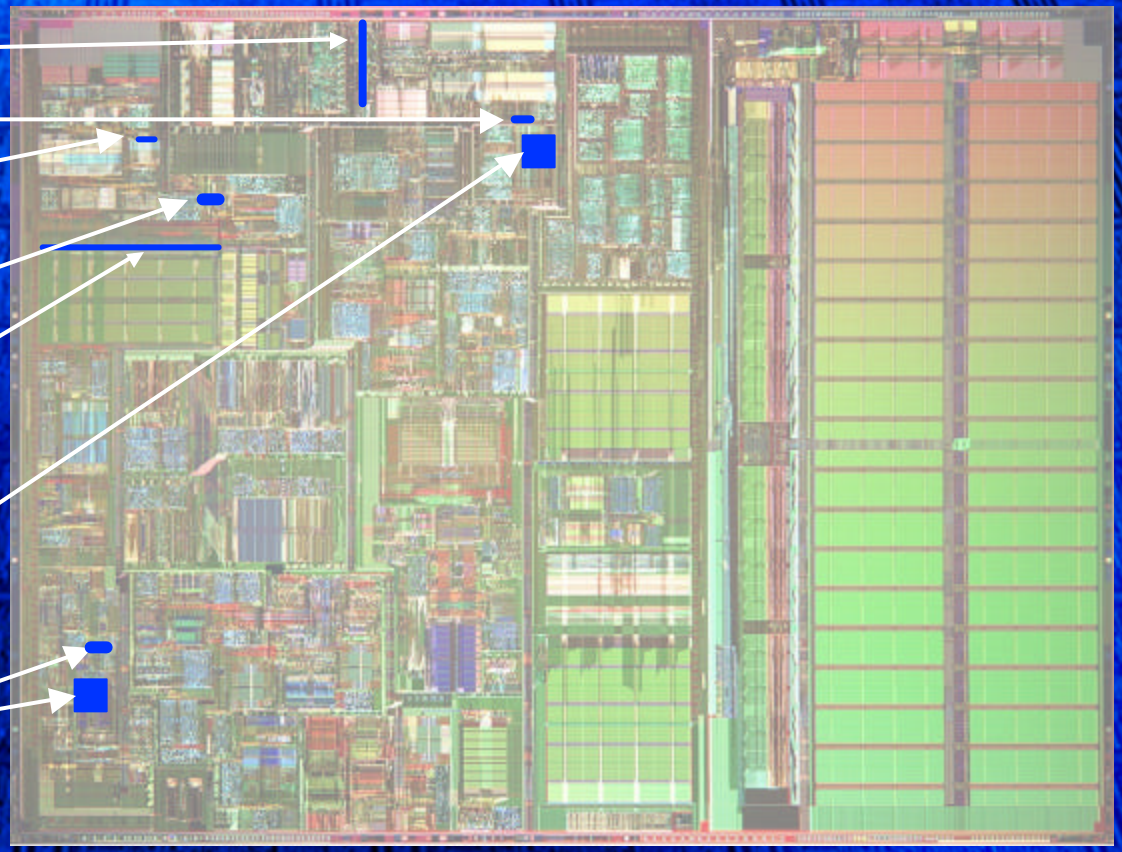



## Hyper-Threading Technology

- Single CPU μArch augmented to look as 2+ CPUs to software
- Improve utilization of processor hardware across high/low ILP code
- Addresses increasing overhead of cache misses
- Speculative multi-threading can improve single-thread performance

# HyperThreading Technology: What was added?

- Instruction Streaming Buffers
- Next Instruction Pointer
- Return Stack Predictor
- Trace Cache Next IP
- Trace Cache Fill Buffers
- Instruction TLB
- Register Alias Tables



~5% die size (& max power), up to 25% performance increase

# Technology Trends

## More Performance

- Multi-Core
- Larger Cache
- Thread-Level Parallelism

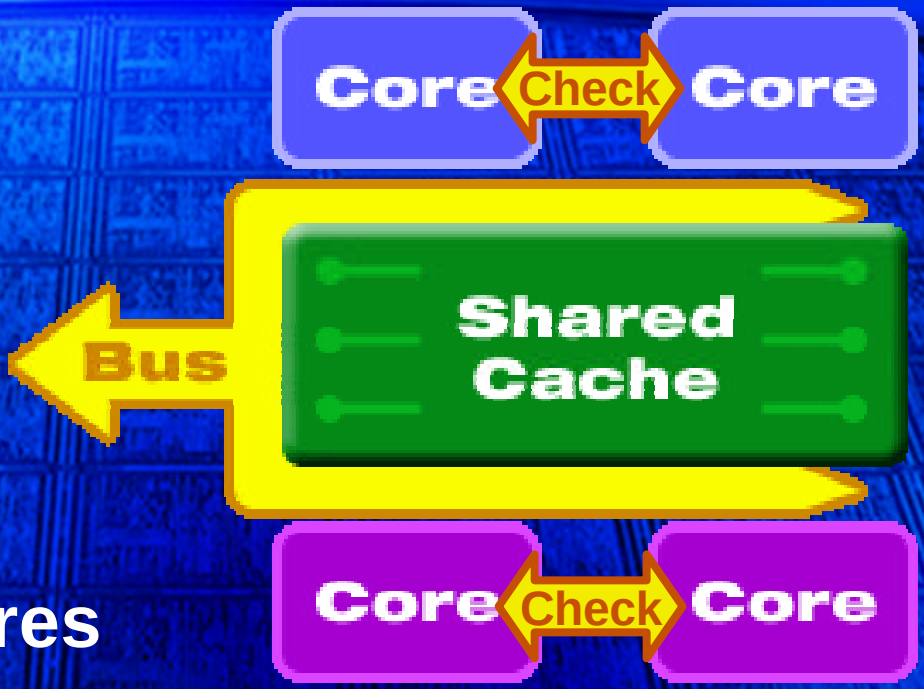
## More Than Performance

- High Reliability
- Managed Run-Time Environments
- Safer Computing

## Balancing Performance and Power

- Power Management
- Design Tradeoffs
- Die Size

# High Reliability Configuration




**FRC Pair the Cores**

- **High Reliability**
- **Shared Cache**
  - **Outside FRC domain**
  - **ECC plus redundant logic**

**1B Transistor Budget Means New Options**

# Managed Run-Time Environments

- Industry Analysts: >50% of Server and Client apps to be .NET* or Java* by 2005
- Rapid Adoption driven by:
  - Developer productivity
  - Better time to market
- Processor Design Impacts:
  - Dynamic optimization/JIT (re-)optimization
  - Dynamic memory management (Garbage Collection)
  - Overhead of modularity, indirect/dynamic specification

**Trade performance for productivity benefits**

# Safer Computing


- Convergence requires a more secure environment
  - HW based strengthening critical
- Intel's LaGrande Technology - a foundation
  - Protected execution, protected memory, protected storage
  - Delivered through processor and chipset extensions
- Intel is working with the industry to enable & deploy

**Intel LaGrande Technology – Basis for Safer Computing**

intel

# Technology Trends

## More Performance

- Multi-Core
- Larger Cache
- Thread-Level Parallelism

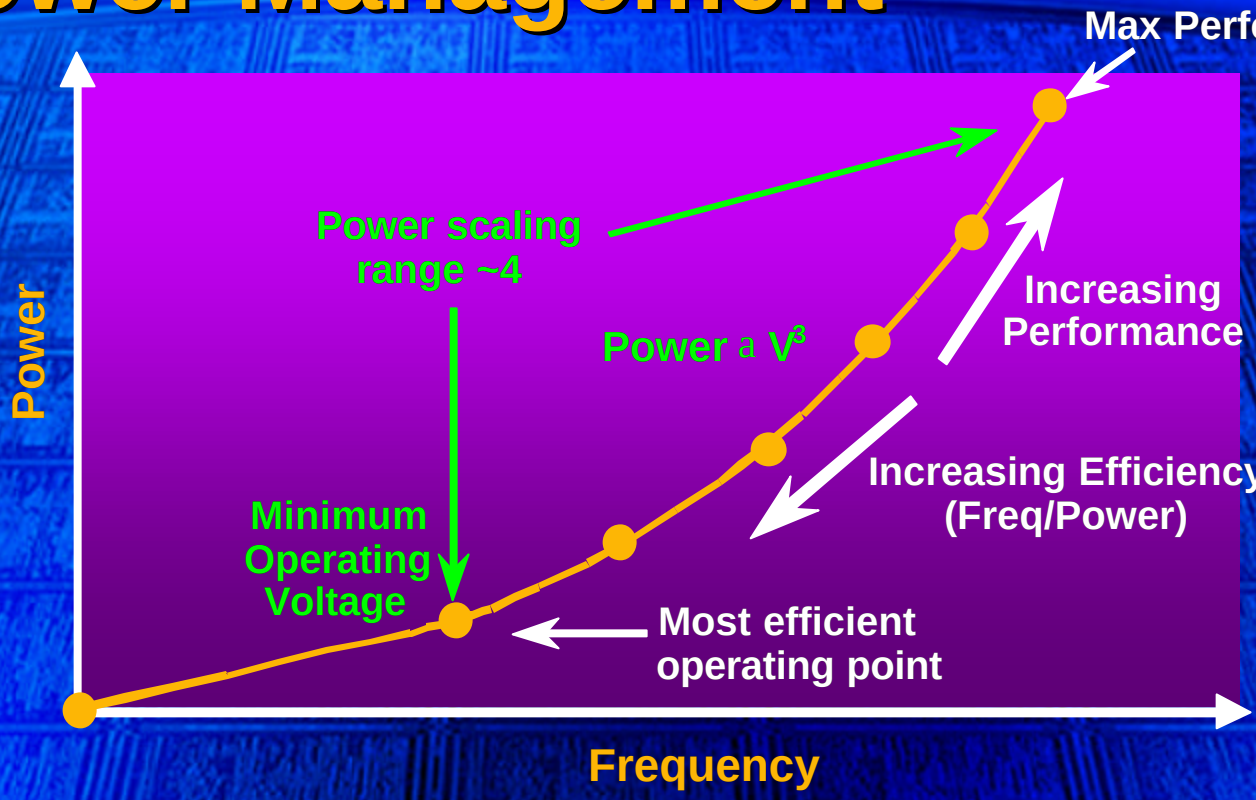
## More Than Performance

- High Reliability
- Managed Run-Time Environments
- Safer Computing

## Balancing Performance and Power

- Power Management
- Design Tradeoffs
- Die Size

# Power Management

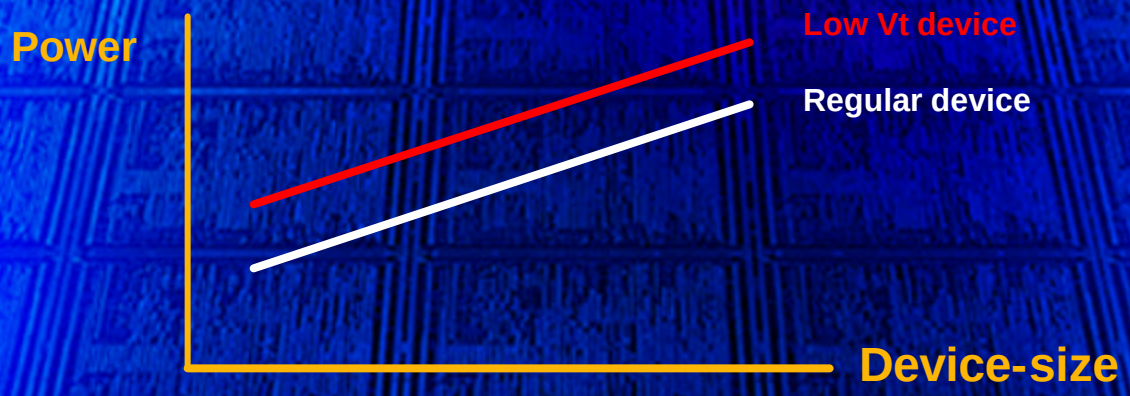



- Enhanced Intel® SpeedStep™ Technology
  - Voltage-frequency scaling with active thermal feedback
  - Operating states range from high perf. to power efficient

**Manage average and peak power dissipation**

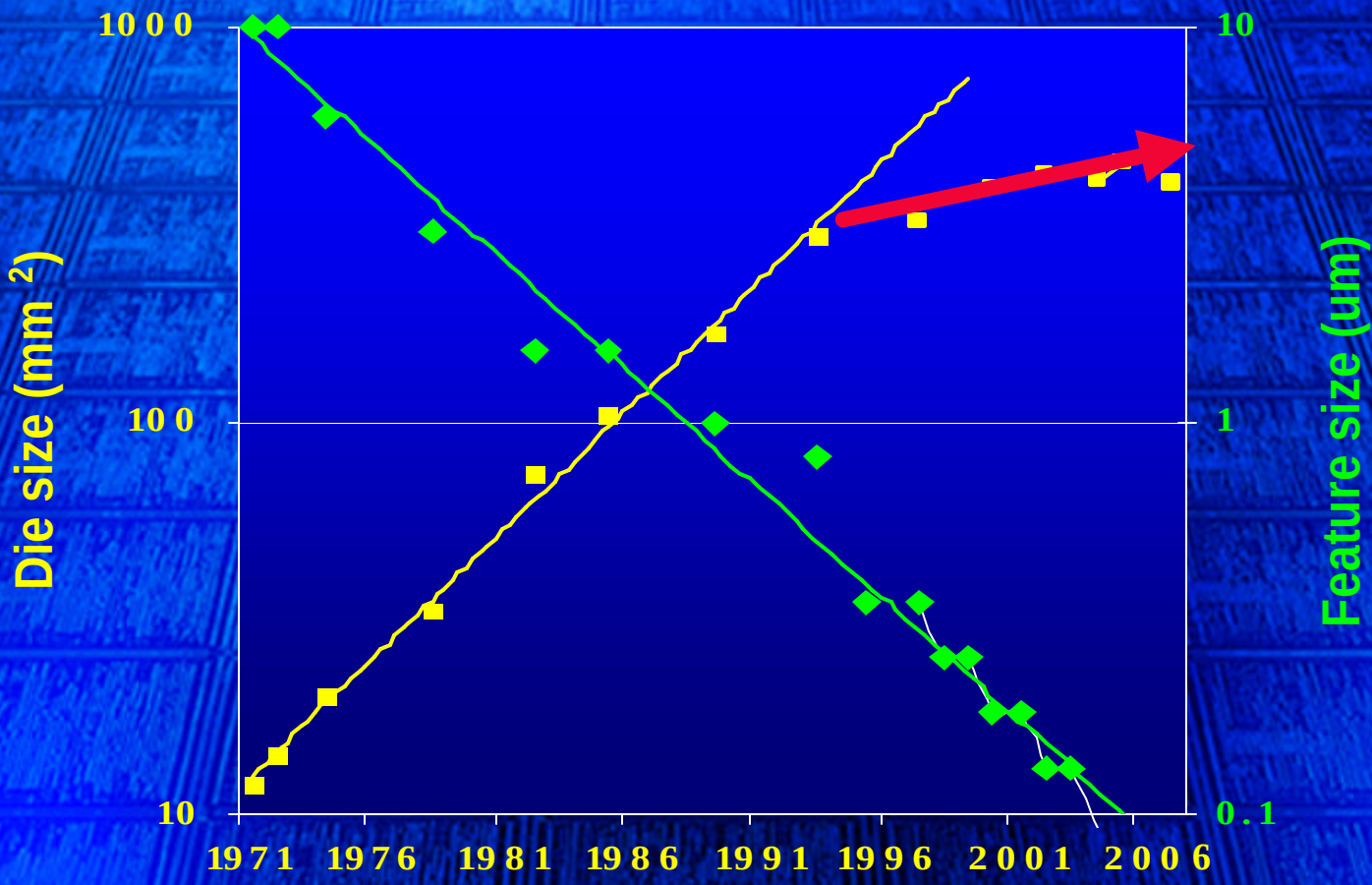
# Full-Chip Timing Optimization

- Goal: Balance signal propagation through logic paths
- Traditional Approach: Speed up devices on critical paths
  - Make devices bigger – faster, but higher active power & leakage
  - Use Lower Vt device – faster, but much higher leakage power
- New Approach: Slow down paths with “too much” slack
  - Reduce device sizes to reduce power and slow down
  - Bring slack paths closer to target period




**Optimize for Performance and Power**

# Feature, Die Size Trend




**Die size growth will be limited by power dissipation**

# 1 Billion Transistors…

- Continued Path of Moore's Law gets us to 1 Billion transistors around 2007
- Performance structures will use 1B transistors
  - Example: Large cache, hyper-threaded multi-core
    - Chip integration of today's board level integration
    - Complexity and power manageable; Addresses performance
- More than Performance: Features address Markets
- Balance Performance and Power

**Major innovations in both performance and beyond as the industry reaches 1 billion transistors**



**http://www.intel.com/pressroom/kits/events/mpf_2002/index.htm**